Attorney Docket No. 50277-1533
(OID #1999-159-01)

PATENT

UNITED STATES PATENT APPLICATION

FOR

DYNAMIC QUALITY ADJUSTMENT BASED ON CHANGING STREAMING CONSTRAINTS

INVENTOR:

DAVID J. PAWSON

PREPARED BY:

HICKMAN PALERMO TRUONG & BECKER LLP
1600 WILLOW STREET
SAN JOSE, CA 95125-5106
(408) 414-1080

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number EL624353281US

Date of Deposit September ~~August~~ 1, 2000

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Casey Moore
(Typed or printed name of person mailing paper or fee)

Casey Moore
(Signature of person mailing paper or fee)

DYNAMIC QUALITY ADJUSTMENT BASED ON CHANGING STREAMING CONSTRAINTS

RELATED APPLICATION DATA

This application is a continuation-in-part application of copending U.S. application Serial No. ~~09/128,224~~ 09/128,244 filed on August 3, 1998, which is a continuation-in-part application of copending U.S. application Serial No. 08/859,860 filed on May 21, 1997, which is a continuation application of U.S. application Serial No. 08/502,480 filed on July 14, 1995, now U.S. Patent No. 5,659,539, all of which are incorporated herein by reference in their entirety.

FIELD OF THE INVENTION

The present invention relates to a method and apparatus for processing audio-visual information, and more specifically, to a method and apparatus for providing improved quality digital media in response to relaxed streaming constraints.

BACKGROUND OF THE INVENTION

In recent years, the media industry has expanded its horizons beyond traditional analog technologies. Audio, photographs, and even feature films are now being recorded or converted into digital formats. Digital media's increasing presence in today's society is not without warrant, as it provides numerous advantageous over analog film. As users of the popular DVD format well know, digital media does not degrade from repeated use. Digital media can also either be delivered for presentation all at once, as when leaded by a DVD player, or delivered in a stream as needed by a digital media server.

As would be expected, the viewers of digital media desire at least the same functionality from the providers of digital media as they now enjoy while watching analog video tapes on video cassette recorders. For example, a viewer of a digital media presentation may wish to mute the audio just as one might in using analog videotapes and

videocassette recorders. Currently, this is performed by adjusting the viewer's volume controls. However, as the server is unaware that audio information is not desired by the viewer, the server still continues to transmit audio information to the viewer. In a distributed digital media environment, the resulting waste in available bandwidth on the digital media server is considerable.

SUMMARY OF THE INVENTION

Techniques are provided for eliminating the waste in bandwidth on the digital media server when a particular type of data is not desired to be received by a user. Extra value is provided to a viewer by utilizing the bandwidth previously allocated to the client to send improved quality images or additional information, such as closed-captioned information. According to one aspect of the present invention, a digital media stream is sent to a client according to a set of streaming constraints. In one embodiment, the digital media stream contains both audio and visual information. According another embodiment, the digital media stream contains only visual information and a separate audio stream is sent to the client containing audio information. Next, a signal is received indicating a relaxation of streaming constraints corresponding to a particular type of data in the digital media stream. In one embodiment, the signal indicates the client is not to receive audio information. In another embodiment, the signal indicates the client is not to receive information of a particular type. In response to the signal, a set of improved quality media information is sent to the client.

According to one embodiment, a set of improved quality media information may be sent using the freed-up portion of the bandwidth previously allocated to the client. According to another embodiment, a set of improved quality media information may be sent to a first client using the freed-up portion of the bandwidth previously allocated to a second client. According to a further embodiment, the set of improved quality media information includes closed-captioned information.

As a result of the techniques described herein, an improved quality digital media stream is available for presentation to a client and, consequently, when a viewer requests to discontinue an undesired component of a streaming video presentation, the undesired information is not sent to the client, which thereby reduces the streaming constraints on a video streaming service, and the improved quality media information may be sent using the freed-up portion of the bandwidth previously allocated to the requesting client.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

Figure 1 is a block diagram of an audio-visual information delivery system according an embodiment of the present invention;

Figure 2 illustrates the various layers of a digital media file according to one embodiment of the present invention;

Figure 3 illustrates the operation of a multiplexor according to an embodiment of the invention;

Figure 4 is a flow chart illustrating the steps of dynamic quality adjustment according to an embodiment of the invention; and

Figure 5 illustrates the operation of a modified multiplexor according to an embodiment of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

A method and apparatus for dynamic quality adjustment based on changing streaming constraints is described. In the following description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form in order to avoid unnecessarily obscuring the present invention.

In the following description, the various features of the invention shall be discussed under topic headings that appear in the following order:

I. SYSTEM OVERVIEW

II. DIGITAL AUDIO/VIDEO FILE STRUCTURE

III. MULTIPLEXOR OPERATIONS

IV. FUNCTIONAL OPERATION

I. SYSTEM OVERVIEW

Figure 1 is a block diagram illustrating an audio-visual information delivery system 100 according to one embodiment of the present invention. Audio-visual information delivery system 100 contains a plurality of clients (1 - n) 160, 170 and 180. The clients (1 - n) 160, 170 and 180 generally represent devices configured to decode audio-visual information contained in a stream of digital audio-visual data. For example, the clients (1 - n) 160, 170 and 180 may be set top converter boxes coupled to an output display, such as a television.

As shown in Figure 1, the audio-visual information delivery system 100 also includes a stream server 110 coupled to a control network 120. Control network 120 may be any network that allows communication between two or more devices. For example, control network 120 may be a high bandwidth network, an X.25 circuit or an electronic industry association (EIA) 232 (RS - 232) serial line or an IP network.

The clients (1- n) 160, 170 and 180, also coupled to the control network 120, communicate with the stream server 110 via the control network 120. For example, clients 160, 170 and 180 may transmit requests to initiate the transmission of audio-visual data streams, transmit control information to affect the playback of ongoing digital audio-visual transmissions, or transmit queries for information. Such queries may include, for example, requests for information about which audio-visual data streams are currently available for service.

The audio-visual information delivery system 100 further includes a video pump 130, a mass storage device 140, and a high bandwidth network 150. The video pump 130 is coupled to the stream server 110 and receives commands from the stream server 110. The video pump 130 is coupled to the mass storage device 140 such that the video pump 130 retrieves data from the mass storage device 140. The mass storage device 140 may be any type of device or devices used to store large amounts of data. For example, the mass storage device 140 may be a magnetic storage device, an optical storage device, or a combination of

such devices. The mass storage device 140 is intended to represent a broad category of non-volatile storage devices used to store digital data, which are well known in the art and will not be described further. While networks 120 and 150 are illustrated as different networks for the purpose of explanation, networks 120 and 150 may be implemented on a single network.

The tasks performed during the real-time transmission of digital media data streams are distributed between the stream server 110 and the video pump 130. Consequently, stream server 110 and video pump 130 may operate in different parts of the network without adversely affecting the efficiency of the system 100.

In addition to communicating with the stream server 110, the clients (1 - n) 160, 170 and 180 receive information from the video pump 130 through the high bandwidth network 150. The high bandwidth network 150 may be any type of circuit-style network link capable of transferring large amounts of data, such as an IP network.

The audio-visual information delivery system 100 of the present invention permits a server, such as the video pump 130, to transfer large amounts of data from the mass storage device 140 over the high bandwidth network 150 to the clients (1 - n) 160, 170 and 180 with minimal overhead. In addition, the audio-visual information delivery system 100 permits the clients (1 - n) 160, 170 and 180 to transmit requests to the stream server 110 using a standard network protocol via the control network 120. In one embodiment, the underlying protocol for the high bandwidth network 150 and the control network 120 is the same. The stream server 110 may consist of a single computer system, or may consist of a plurality of computing devices configured as servers. Similarly, the video pump 130 may consist of a single server device, or may include a plurality of such servers.

To receive a digital audio-visual data stream from a particular digital audio-visual file, a client (1 - n) 160, 170 or 180 transmits a request to the stream server 110. In response to the request, the stream server 110 transmits commands to the video pump 130 to cause video pump 130 to transmit the requested digital audio-visual data stream to the client that

requested the digital audio-visual data stream.

The commands sent to the video pump 130 from the stream server 110 include control information specific to the client request. For example, the control information identifies the desired digital audio-visual file, the beginning offset of the desired data within the digital audio-visual file, and the address of the client. In order to create a valid digital audio-visual stream at the specified offset, the stream server 110 may also send "prefix data" to the video pump 130 and may request the video pump 130 to send the prefix data to the client. Prefix data is data that prepares the client to receive digital audio-visual data from the specified location in the digital audio-visual file.

The video pump 130, after receiving the commands and control information from the stream server 110, begins to retrieve digital audio-visual data from the specified location in the specified digital audio-visual file on the mass storage device 140.

The video pump 130 transmits any prefix data to the client, and then seamlessly transmits digital audio-visual data retrieved from the mass storage device 140 beginning at the specified location to the client via the high bandwidth network 150.

The requesting client receives the digital audio-visual data stream, beginning with any prefix data. The client decodes the digital audio-visual data stream to reproduce the encoded audio-visual sequence.

II. DIGITAL AUDIO/VIDEO FILE STRUCTURE

Having described the system overview of the audio-visual information delivery system 100, the format of the digital media, or audio-visual, file structure will now be described. Digital audio-visual storage formats, whether compressed or not, use state machines and packets of various structures. The techniques described herein apply to all such storage formats. While the present invention is not limited to any particular digital audio-visual format, the MPEG-2 transport file structure shall be described for the purposes of illustration.

Referring to Figure 2, it illustrates the structure of an MPEG-2 transport file 104 in greater detail. The data within MPEG file 104 is packaged into three layers: a program elementary stream ("PES") layer, a transport layer, and a video layer. These layers are described in detail in the MPEG-2 specifications. At the PES layer, MPEG file 104 consists of a sequence of PES packets. At the transport layer, the MPEG file 104 consists of a sequence of transport packets. At the video layer, MPEG file 104 consists of a sequence of picture packets. Each picture packet contains the data for one frame of video.

Each PES packet has a header that identifies the length and contents of the PES packet. In the illustrated example, a PES packet 250 contains a header 248 followed by a sequence of transport packets 251-262. PES packet boundaries coincide with valid transport packet boundaries. Each transport packet contains exclusively one type of data. In the illustrated example, transport packets 251, 256, 258, 259, 260 and 262 contain video data. Transport packets 252, 257 and 261 contain audio data. Transport packet 253 contains control data. Transport packet 254 contains timing data. Transport packet 255 is a padding packet.

Each transport packet has a header. The header includes a program ID ("PID") for the packet. Packets assigned PID 0 are control packets. For example, packet 253 may be assigned PID 0. Control packets contain information indicative of what programs are present in the digital audio-visual data stream. Control packets associate each program with the PID numbers of one or more PMT packets, which contain Program Map Tables. Program Map Tables indicate what data types are present in a program, and the PID numbers of the packets that carry each data type. Illustrative examples of what data types may be identified in PMT packets include, but are not limited to, MPEG2 video, MPEG2 audio in English, and MPEG2 audio in French.

In the video layer, the MPEG file 104 is divided according to the boundaries of frame data. As mentioned above, there is no correlation between the boundaries of the data that represent video frames and the transport packet boundaries. In the illustrated example, the

frame data for one video frame "F" is located as indicated by brackets 270. Specifically, the frame data for frame "F" is located from a point 280 within video packet 251 to the end of video packet 251, in video packet 256, and from the beginning of video packet 258 to a point 282 within video packet 258. Therefore, points 280 and 282 represent the boundaries for the picture packet for frame "F". The frame data for a second video frame "G" is located as indicated by brackets 272. The boundaries for the picture packet for frame "G" are indicated by bracket 276.

Many structures analogous to those described above for MPEG-2 transport streams also exist in other digital audio-visual storage formats, such as MPEG-1, Quicktime, and AVI. In one embodiment, indicators of video access points, time stamps, file locations, etc. are stored such that multiple digital audio-visual storage formats can be accessed by the same server to simultaneously serve different clients from a wide variety of storage formats. Preferably, all of the format specific information and techniques are incorporated in the stream server. All of the other elements of the server are format independent.

III. MULTIPLEXOR OPERATIONS

It is often desirable to merge several digital media presentations, each presentation in a separate digital media stream, into one stream containing the combined digital media presentations. This merger allows a user to select different digital media presentations to watch from a single digital media stream. Figure 3 illustrates a multiplexor 310, which is a digital media component that performs the operation of merging multiple digital media streams into a single digital media stream. As multiplexors are well understood to those in the art, description in this section will be limited to the extent that it facilitates understanding of their use in optimizing mute operations in a multiplexed stream environment, which will be described in detail below.

As Figure 3 shows, a multiplexor 310 has multiple inputs and a single output. The inputs to the multiplexor are called Single Program Transport Streams ("SPTS"), labeled as

320, 322, and 324, and the output is called a Multiple Program Transport Stream ("MPTS"), which is labeled as 330. A Single Program Transport Stream 320, 322, and 324 is a digital media stream that is encoded with audio and video data for one video presentation. Alternately, a Multiple Program Transport Stream 330 is a digital media stream that is encoded with audio and video data for multiple video presentations. Thus, a Single Program Transport Stream 320, 322, and 324 is analogous to a single channel on TV, whereas a Multiple Program Transport Stream 330 is analogous to a cable network.

When the individual SPTSs 320, 322, and 324 are combined, the multiplexor 310 examines the PID in each transport packet to ensure that each PID referenced in the control packets is unique. In the case when packets from different SPTSs 320, 322, and 324 use the same PID, the multiplexor 310 remaps the PIDs to unique numbers to ensure that each packet can easily be identified as belonging to a particular Single Program Transport Stream 320, 322, and 324. As each audio and video packet is guaranteed to have a unique PID, the video presentation to which the packet corresponds may be easily identified by examining the PID 0 control packets in the MPTS 330. Thus, as the multiplexor 310 must examine each table in the PID 0 control packets and all tables of packets references in the PID 0 control packets to ensure all referenced packets have a unique PID number, it also can easily identify all audio packets corresponding to a particular SPTS 320, 322, and 324.

IV. FUNCTIONAL OPERATION

A client may reduce the amount of a particular type of information contained in the digital media presentation that is received. In one embodiment, the amount of a particular type of information required by the client is reduced as the result of altering the presentation characteristics to a state requiring less of the particular type of information, such as when reducing the video resolution, or switching the sound output from stereo to mono. In another embodiment, the particular type of information is not required at all, such as when a client mutes the audio portion of a presentation. It is beneficial for the stream server 110 to reclaim

the bandwidth previously allocated to delivering that particular type of information to the client. This extra bandwidth can be used to improve the quality of the digital media presentation, or to send additional information, such as closed-captioned information.

An exemplary description will now be provided with reference to Figure 4 to illustrate the process of reclaiming unused bandwidth wherein the client mutes the audio in a digital media presentation. The client 160 sends a signal through the control network 120 to the stream server 110 to indicate that audio data is not to be sent to the client. The signal is sent using existing communication protocols, such as Real Time Streaming Protocol ("RTSP").

In one embodiment, the stream server 110 operates in a multiplexed environment, or an environment in which audio and visual data is sent to the client in a single stream, such as in MPEG. In response to receiving the signal, a multiplexor is used to examine and identify the packets for the particular SPTS being muted. The multiplexor then discards the identified audio packets for the muted SPTS and does not combine them in the output stream.

In another embodiment, the stream server 110 still operates in a multiplexed environment, but in response to receiving the signal, a modified multiplexor 510 is used to examine and identify the packets for the particular SPTS being muted, as shown in Figure 5. The modified multiplexor 510 operates in substantially the same way as described in the prior section, except that it operates with only one input SPTS 520. The modified multiplexor 510 then filters and discards the identified audio packets for the input SPTS 520. The resulting output stream 530 from the modified multiplexor 510 contains the original media presentation, but not any audio packets, from the input SPTS 520.

In still another embodiment, the stream server 110 operates in a split-stream environment, or an environment in which audio and visual data are sent to the client in separate streams. In response to receiving the signal, the stream server 110 continues sending the video stream, but pauses or stops sending the audio stream to the signaling client. As the video is sent in a different stream to the signaling client than the audio, stopping the

audio stream will not interrupt the video presentation to the signaling client.

As audio packets for the muted digital video stream are no longer sent to the client, the bandwidth previously allocated to the signaling client can be reclaimed. Accordingly, streaming constraints on the stream server 110 are reduced.

As mentioned previously, reclaiming bandwidth as a result of a client signaling to discontinue transmission of a particular type of information is not limited to audio information. A client may signal to indicate any particular type of information contained within the digital media stream is no longer to be sent to that client. For example, the client signals to indicate that visual information in no longer to be sent. Accordingly, the reclaimed bandwidth on the stream server 110 may be used to send improved quality information of the remaining types of information contained in the digital media stream, or send additional information. For example, if a client signals to indicate visual information is not to be sent, improved quality audio information may be sent. Examples of improved quality audio information include, but are not limited to, sending audio information in a format such as THX or Dolby, sending additional sound tracks, or sending information in surround sound.

In one embodiment, bandwidth reclaimed on the stream server 110 from one client may be utilized by any client of the stream server 110. In another embodiment, bandwidth reclaimed on the stream server 110 from one client may only be used by that client.

As mentioned above, one use of the reclaimed bandwidth is to provide improved quality. The quality of the video may be improved by modifying one or more of a video's characteristics. Examples of improving the quality of a video include, but are not limited to, increasing the rate of frame transmission, increasing color depth, and increasing the pixel density. In addition to, or instead of, increasing the quality of the video, the reclaimed bandwidth may be used to send or improve other data associated with the video. For example, the reclaimed bandwidth may be used to send closed-captioned information, additional information, or otherwise alter the appearance of the video in some form.

In other embodiments, the quality of the video may be improved through improved

quantization. Improved quantization is achieved by collapsing similar states into a single state, thereby allowing more unique states to be identified. For example, assume each color used in a digital video presentation is assigned a 24 bit number. By grouping similar colors together and assigning them the same 24 bit number, more unique colors may be identified for use in the digital video with 24 bits.

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.